The chromosomal location of genes for elongation factor Tu and ribosomal protein S10 in the cyanobacterium *Spirulina platensis* provides clues to the ancestral organization of the *str* and S10 operons in prokaryotes

**Anna M. Sanangelantoni* and Orsola Tiboni**

Dipartimento di Genetica e Microbiologia 'A. Buzzati-Traverso', Università di Pavia, via Abbiategrasso 207, 27100 Pavia, Italy

(Received 2 March 1993; revised 9 June 1993; accepted 14 June 1993)

---

Introduction

The organization and regulation of genes encoding components of the translational apparatus have been extensively investigated in *Escherichia coli* (for a review, see Lindahl & Zengel, 1986). As more data become available in other prokaryotes, it appears that the chromosomal organization of these genes is substantially conserved amongst bacteria and archaea, although differences do exist in detail.

In the cyanobacteria *Spirulina platensis* and *Anacystis nidulans* the genes encoding ribosomal proteins S12, S7 and elongation factors EF-G (*fus*) and EF-Tu (*tuf*) are arranged in a transcriptional unit resembling the *str* operon of *Escherichia coli* (Buttarelli et al., 1989; Meng et al., 1989). An *E. coli*-like organization of these genes is also maintained in the cyanelles of *Cyanophora paradoxa* (for a review, see Lindahl & Zengel, 1986). As more data become available in other prokaryotes, it appears that the chromosomal organization of these genes is substantially conserved amongst bacteria and archaea, although differences do exist in detail.

The nucleotide sequence data reported in this paper have been submitted to the EMBL Data Library and assigned the accession number Z21676.

---

* Author for correspondence. Tel. +39 382 391578; fax +39 382 528496.

**Abbreviation**: EF, elongation factor.
short distance (only 50 bp) compared to that (15 kb) between the \textit{tuf} and \textit{rps10} genes in \textit{E. coli} (O. Tiboni and others, unpublished data). Given that the linkage between the \textit{tuf} and \textit{rps10} genes is maintained in cyanelles, in all of the archaea, in \textit{Tg. maritima} and, as shown in this paper, in the cyanobacterium \textit{Spirulina platensis}, the possibility exists that linkage of the \textit{rps10} gene with those of the \textit{str} operon represents the ancestral situation.

Here we report the cloning, sequencing and localization of the \textit{rps10} gene relative to the \textit{tuf} gene in \textit{S. platensis}.

\section*{Methods}

\textbf{Bacterial strains and plasmids.} \textit{Spirulina platensis} cells were grown as reported by Riccardi et al. (1981). A plasmid containing a 1250 bp \textit{EcoRI-PstI} fragment of \textit{Cyanophora paradoxa} cyanellar DNA (kindly supplied by Dr W. Loffelhardt) in the vector pMB18 was used to prepare the heterologous probe. Plasmid pSP7 (Tiboni & Di Pasquale, 1987), containing a 52 kb fragment of \textit{S. platensis} genomic DNA, was used for subcloning. Plasmids pBR322 (Bolivar \textit{et al.}, 1977) and pUC19 (Vieira & Messing, 1982) were used as vectors and \textit{E. coli} HB101 as the host. Plasmid-containing strains were grown in LB medium supplemented with 50 \mu g ampicillin ml$^{-1}$.

\textbf{Recombinant DNA techniques.} Chromosomal DNA from \textit{S. platensis} was prepared as described previously (Tiboni \textit{et al.}, 1984). Plasmid DNA was isolated and purified using Qiagen tips (Diagen) following the manufacturer's instructions. Recovery of DNA fragments from low-melting-point agarose gel was performed as reported by Higuchi \textit{et al.} (1985). Transformation experiments were carried out as described by Dagtger & Ehrlich (1979). Hybridization experiments (Southern, 1975) were performed at 37 $^\circ$C in the presence of 55\% (v/v) formamide (for heterologous DNA) or at 42 $^\circ$C in the presence of 50\% formamide (homologous DNA). Probes were labelled by the Multiprime DNA Labelling System using [$\alpha$-\textit{32P}]dCTP [Amersham, sp. act. 3000 Ci mmol$^{-1}$ (111 TBq)] and the Multiprime DNA Labelling System kit supplied by Amersham. Restriction endonuclease enzymes and T4 DNA ligase were obtained from Boehringer and used according to the manufacturer's instructions.

\textbf{Sequencing.} The inserts of plasmids pSP14 and pSP7 were digested with several restriction enzymes and fragments of 800–1000 bp were subcloned in the polylinker cloning sites of pUC19 and sequenced by the dideoxy chain-termination method of Sanger \textit{et al.} (1977). In the case of longer inserts, synthetic oligonucleotides were used as primer. The DNA was radiolabelled with [$\alpha$-\textit{32P}]dATP [\textit{greater than} 1000 Ci mmol$^{-1}$ (37 TBq)] and primed as specified by the manufacturer using a T7 Sequencing kit (Pharmacia).

Analysis of the DNA sequence was performed using a Macintosh IIcx computer and the DNA Strider 1.1 program (Marck, 1988).

\textbf{Phylogenetic analysis.} Amino acid sequences were aligned by the hierarchical clustering method of Corpet (1988) as implemented in the program \textsc{multalin} with a gap penalty of 8. The amino acid sequence alignment was converted into a colinear alignment of first and second codon positions and phylogenetic trees were constructed by distance matrix and maximum parsimony methods from clearly homologous gap-free regions of the DNA alignments using the programs \textsc{consense}, \textsc{fitch}, \textsc{dnadist}, \textsc{dnapars} and \textsc{seqboot} as implemented in the \textsc{phylip} 3.5e package. Evolutionary distances were calculated from the pairwise nucleotide sequence identities using the Jukes–Cantor model of sequence evolution as implemented in the \textsc{dnadist} program (Felsenstein, 1990). The reliability of the internal branches was determined by bootstrap analysis of at least 100 pseudoreplications of the original sample.

\section*{Results and Discussion}

\textbf{Cloning of the \textit{S. platensis} \textit{rps10} gene}

To identify DNA fragments carrying the \textit{rps10} gene of \textit{S. platensis}, genomic DNA was digested with several restriction endonucleases and probed with the 1-25 kb \textit{EcoRI-PstI} fragment of \textit{C. paradoxa} cyanellar DNA containing the entire coding region of the \textit{rps10} gene as well as the entire \textit{petFL} gene and the 3' terminal region of the \textit{tuf} gene (Neumann-Spallart \textit{et al.}, 1991; Bryant \textit{et al.}, 1991). Among all digestions exhibiting different positive signals, the \textit{HaeIII} digestion was selected because the size of the two fragments (1800 and 600 bp) hybridizing to the probe allowed discrimination between the signal due to the previously cloned \textit{tuf} gene (600 bp fragment, Buttarelli \textit{et al.}, 1989) and that due to \textit{rps10} and/or \textit{petFL} genes. Thus, a partial genomic library of \textit{S. platensis} DNA (1–6–20 kb \textit{HaeIII} fragments) was constructed in \textit{E. coli} HB101 in the \textit{EcoRV} site of pBR322.

Since preliminary experiments demonstrated that the signal produced by the \textit{S. platensis} \textit{rps10} gene was as weak as that produced by the gene of the host, it was impossible to carry out colony hybridizations. Thus, the recombinant plasmids carrying sequences homologous to the probe were selected by analysis of the inserts. Plasmid DNA was prepared from 30 pools of 18 recombinant clones each, and digested with \textit{HaeIII}. The fragments were separated by gel electrophoresis and hybridized with the cyanellar probe. The individual transformants from the positive pool were analysed for plasmid content and a recombinant plasmid, containing a 1-8 kb \textit{HaeIII} fragment (designated pSP14) was isolated. A preliminary sequence analysis indicated that the 3' end of the insert contained part of the \textit{S. platensis} \textit{rps10} gene.

Because in \textit{C. paradoxa} cyanelles the \textit{rps10} gene is located 87 bp downstream from the stop codon of the \textit{tuf} gene (Neumann-Spallart \textit{et al.}, 1991), experiments were performed to ascertain if the same organization was present in the phylogenetically related cyanobacteria. To this end, the 5'2 kb fragment of the plasmid pSP7 (Tiboni & Di Pasquale, 1987) harbouring part of the \textit{S. platensis} \textit{str} operon (namely, part of the \textit{fus} gene and the entire \textit{tuf} gene) and a downstream region of 2200 bp, was probed with the insert of pSP14 demonstrating that the entire \textit{HaeIII}/\textit{HaeIII} fragment cloned in plasmid pSP14 was contained within this region. Thus, starting from the 3' end of the \textit{S. platensis} \textit{tuf} gene, the nucleotide sequence of 1215 bp was determined. In addition, and because in
bacteria (E. coli and Mycoplasma capricolum) the rps10 gene is the promoter distal gene of the S10 operon, the sequence was extended to cover the 3' end of the pSp7 insert using synthetic 15-mer oligonucleotides (data not shown).

**Nucleotide sequence of the S. platensis S10 structural gene**

The sequence of the 1215 bp region, whose relevant part is shown in Fig. 1, revealed the presence of an ORF of 318 bp located 143 bp downstream of the tuf gene. The derived amino acid sequence shared a high degree of similarity with the cyanellar ribosomal protein S10. The length of the intergenic region between the tuf and rps10 genes was comparable to that separating the genes in the S. platensis str operon. In addition, no obvious p-independent termination structures, could be identified within the intergenic region, suggesting the possibility that rps10 and tuf genes are cotranscribed as in the case of the cyanelles of C. paradoxa.

**Amino acid sequence comparison of S. platensis ribosomal protein S10 with those of cyanelles, archaea and bacteria**

The deduced amino acid sequence of the S. platensis ribosomal protein S10 was compared with available homologues from bacteria (E. coli, Tg. maritima, Myco. capricolum), archaea (Sulfolobus acidocaldarius, Methanococcus vannielii, Pyrococcus woesei, Thermoplasma acidophilum), euakarya (rat) and cyanelles of C. paradoxa. No other plastidial S10 sequences were available for comparison. The alignment presented in Fig. 2 clearly shows that all of the S10-equivalent sequences contain the conserved heptapeptide GPIPLPT that could be related to the functioning of S10 in the ribosome. S10 is, in fact, located on the external surface of the small subunit of the E. coli ribosome together with proteins S3, S5, S14 and S19, all of which appear to be involved in tRNA binding (Kahan et al., 1981). As the conserved heptapeptide occurs in all of the prokaryotic S10-equivalent proteins examined until now, it is reasonable...
parsimony trees show a clear segregation of the cyanellar (Cya), bacteria (S. platensis, Spl; E. coli, Eco; Myco. capricolum, Mca; Tg. maritima, Tma), archaea (Meth. vannielli, Mva; Tpl. acidophilum, Tac; Sulf. acidocaldarius, Sac; P. woesei, Pwo) and eukarya (Rat).

Table 1. Identity matrices from nucleotide and amino acid sequence alignments of ribosomal protein S10

<table>
<thead>
<tr>
<th></th>
<th>Spl</th>
<th>Cya</th>
<th>Eco</th>
<th>Mca</th>
<th>Tma</th>
<th>Mva</th>
<th>Tac</th>
<th>Sac</th>
<th>Pwo</th>
<th>Rat</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spl</td>
<td>79.5</td>
<td>73.5</td>
<td>65.7</td>
<td>70.6</td>
<td>47.8</td>
<td>58.5</td>
<td>51.5</td>
<td>49.5</td>
<td>50.5</td>
<td></td>
</tr>
<tr>
<td>Cya</td>
<td>74.0</td>
<td>65.7</td>
<td>63.7</td>
<td>69.6</td>
<td>50.0</td>
<td>55.0</td>
<td>54.0</td>
<td>50.0</td>
<td>52.4</td>
<td></td>
</tr>
<tr>
<td>Eco</td>
<td>59.4</td>
<td>52.1</td>
<td>66.8</td>
<td>67.5</td>
<td>47.7</td>
<td>59.3</td>
<td>52.1</td>
<td>50.5</td>
<td>50.0</td>
<td></td>
</tr>
<tr>
<td>Mca</td>
<td>50.0</td>
<td>47.9</td>
<td>51.0</td>
<td>68.8</td>
<td>52.3</td>
<td>60.3</td>
<td>56.2</td>
<td>56.2</td>
<td>50.5</td>
<td></td>
</tr>
<tr>
<td>Tma</td>
<td>64.6</td>
<td>61.5</td>
<td>55.2</td>
<td>56.3</td>
<td>52.8</td>
<td>58.2</td>
<td>55.6</td>
<td>55.7</td>
<td>51.0</td>
<td></td>
</tr>
<tr>
<td>Mva</td>
<td>34.1</td>
<td>31.8</td>
<td>32.9</td>
<td>35.3</td>
<td>35.3</td>
<td>70.3</td>
<td>65.9</td>
<td>76.9</td>
<td>59.3</td>
<td></td>
</tr>
<tr>
<td>Tac</td>
<td>38.9</td>
<td>34.7</td>
<td>38.9</td>
<td>35.8</td>
<td>41.1</td>
<td>54.7</td>
<td>69.3</td>
<td>72.5</td>
<td>59.9</td>
<td></td>
</tr>
<tr>
<td>Sac</td>
<td>35.8</td>
<td>34.7</td>
<td>32.6</td>
<td>30.5</td>
<td>35.8</td>
<td>44.2</td>
<td>45.8</td>
<td>72.0</td>
<td>53.5</td>
<td></td>
</tr>
<tr>
<td>Pwo</td>
<td>35.8</td>
<td>32.6</td>
<td>36.8</td>
<td>35.8</td>
<td>41.1</td>
<td>67.4</td>
<td>58.4</td>
<td>47.9</td>
<td>59.0</td>
<td></td>
</tr>
<tr>
<td>Rat</td>
<td>24.2</td>
<td>31.6</td>
<td>25.3</td>
<td>24.2</td>
<td>28.4</td>
<td>39.5</td>
<td>38.5</td>
<td>41.7</td>
<td>28.1</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Spl</th>
<th>Cya</th>
<th>Eco</th>
<th>Mca</th>
<th>Tma</th>
<th>Mva</th>
<th>Tac</th>
<th>Sac</th>
<th>Pwo</th>
<th>Rat</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spl</td>
<td>79.5</td>
<td>73.5</td>
<td>65.7</td>
<td>70.6</td>
<td>47.8</td>
<td>58.5</td>
<td>51.5</td>
<td>49.5</td>
<td>50.5</td>
<td></td>
</tr>
<tr>
<td>Cya</td>
<td>74.0</td>
<td>65.7</td>
<td>63.7</td>
<td>69.6</td>
<td>50.0</td>
<td>55.0</td>
<td>54.0</td>
<td>50.0</td>
<td>52.4</td>
<td></td>
</tr>
<tr>
<td>Eco</td>
<td>59.4</td>
<td>52.1</td>
<td>66.8</td>
<td>67.5</td>
<td>47.7</td>
<td>59.3</td>
<td>52.1</td>
<td>50.5</td>
<td>50.0</td>
<td></td>
</tr>
<tr>
<td>Mca</td>
<td>50.0</td>
<td>47.9</td>
<td>51.0</td>
<td>68.8</td>
<td>52.3</td>
<td>60.3</td>
<td>56.2</td>
<td>56.2</td>
<td>50.5</td>
<td></td>
</tr>
<tr>
<td>Tma</td>
<td>64.6</td>
<td>61.5</td>
<td>55.2</td>
<td>56.3</td>
<td>52.8</td>
<td>58.2</td>
<td>55.6</td>
<td>55.7</td>
<td>51.0</td>
<td></td>
</tr>
<tr>
<td>Mva</td>
<td>34.1</td>
<td>31.8</td>
<td>32.9</td>
<td>35.3</td>
<td>35.3</td>
<td>70.3</td>
<td>65.9</td>
<td>76.9</td>
<td>59.3</td>
<td></td>
</tr>
<tr>
<td>Tac</td>
<td>38.9</td>
<td>34.7</td>
<td>38.9</td>
<td>35.8</td>
<td>41.1</td>
<td>54.7</td>
<td>69.3</td>
<td>72.5</td>
<td>59.9</td>
<td></td>
</tr>
<tr>
<td>Sac</td>
<td>35.8</td>
<td>34.7</td>
<td>32.6</td>
<td>30.5</td>
<td>35.8</td>
<td>44.2</td>
<td>45.8</td>
<td>72.0</td>
<td>53.5</td>
<td></td>
</tr>
<tr>
<td>Pwo</td>
<td>35.8</td>
<td>32.6</td>
<td>36.8</td>
<td>35.8</td>
<td>41.1</td>
<td>67.4</td>
<td>58.4</td>
<td>47.9</td>
<td>59.0</td>
<td></td>
</tr>
<tr>
<td>Rat</td>
<td>24.2</td>
<td>31.6</td>
<td>25.3</td>
<td>24.2</td>
<td>28.4</td>
<td>39.5</td>
<td>38.5</td>
<td>41.7</td>
<td>28.1</td>
<td></td>
</tr>
</tbody>
</table>

to argue that it is involved in this step of protein synthesis.

The identity matrices, reported in Table 1, were calculated from gap-free sites of the amino acid sequence alignment in Fig. 2 (lower left quadrant) and from a colinear alignment of the corresponding DNA sequence (first and second codon position; upper right quadrant). As expected the degree of similarity between the S. platensis ribosomal protein S10 and the cyanellar S10 is higher (74% amino acid identity) than the similarity between the Spirulina protein and that of other bacteria. A much lower degree of similarity (about 30% amino acid identity) exists between the cyanobacterial and the archaean S10 protein. The mutual relationships between the S10 sequences are concisely illustrated by the distance tree in Fig. 3(a), constructed from DNA sequence identities using the Jukes-Cantor correction for determination of evolutionary distances, and by the tree in Fig. 3(b) constructed from DNA sequence alignment using the maximum parsimony method. In agreement with results obtained using other phylogenetic probes, such as elongation factors (Auer et al., 1991; Cammarano et al., 1992) and DNA polymerase core subunits (Pühler et al., 1989), both DNA sequence distance and parsimony trees show a clear segregation of S10 sequences from the three primary domains, the archaean sequences being definitely closer to the eukaryal than to the bacterial ones. Also, in both trees cyanobacterial and cyanelles form sister branches at high (95-100%) bootstrap confidence levels, and the overall tree topologies from the two methods are highly congruent, both being insensitive to removal of eukaryal and archaean representatives. The deep branching position of Mycoplasma (Mycoplasma not Thermotaoga being the deepest offshoot in the bacterial branch) most probably reflects a disparity of evolutionary rates due to the reported fast evolution of the former species (Woese, 1987).

Organization of the tuf and rps10 gene

Unlike those of eukaryotes, ribosomal protein genes of archaea and bacteria are arranged in transcriptional units. The organization of those operons has been
Although no information is available concerning the tuf gene, in that order), whilst the rps10 gene is the first unit of another operon (S10 operon) located far downstream (15 kb) from the tuf gene. Although no information is available concerning the tuf gene in Myco. capricolum, the rps10 gene has been cloned and sequenced together with the genes constituting the S10-spc operon (Ohkubo et al., 1989). In P. woesei (Creti et al., 1991) and Sulf. acidocaldarius (Auer et al., 1989); and in H. marismortui a gene cluster equivalent to the S10 operon has been cloned and sequenced but the S10-encoding gene appears to be located elsewhere in the genome (Arndt et al., 1990). In P. woesei (Creti et al., 1991) and Sulf. acidocaldarius (Auer et al., 1991) the S10-encoding gene is linked to the tuf gene and immediately followed by a tRNA^"\*" gene.

This is not the first example of deviation of the cyanobacterial ribosomal protein gene organization from the E. coli pattern. The genes for ribosomal proteins L10 and L12, which in E. coli cluster in the β operon with the rpoB and rpoC genes, appear to be unlinked from the RNA polymerase subunits genes in Synechocystis 6803 (Sibold & Subramanian, 1990). In archaea the same genes are also widely separated (Matheson et al., 1990). Similarity in gene organization between cyanobacteria and archaea has also been detected in the case of the genes encoding the larger subunit of RNA polymerase (Bergsland & Haselkorn, 1991).

Some conclusions can be drawn from the results of this study. Firstly, the strong similarity between the cyanellar and cyanobacterial S10 sequence provides further support for the endosymbiotic origin of cyanelles from a free-living cyanobacterium. Secondly, the close linkage between rps10 and tuf genes suggests that this organization has been retained in the cyanelles. Thirdly, this new case (after rpo genes) of colinearity in gene organization between such distantly related organisms as cyanobacteria and archaea suggests the possibility that this gene arrangement represents the ancestral situation.

This work was supported by grants of the MURST and from a CNR grant Progetto Finalizzato Biotecnologie e Biostrumentazione, sotto-progetto Bioingegneria Molecolare e Cellulare. We are greatly indebted to Dr. Piero Cammarano for his most generous help in deriving phylogenetic trees.
References


